Lock and Key to Transcription: σ-DNA Interaction

Xin Liu,[1] David A. Bushnell,[1] and Roger D. Kornberg[1,*]
[1]Department of Structural Biology, Stanford University School of Medicine, Stanford, CA 94305, USA
*Correspondence: kornberg@stanford.edu
DOI 10.1016/j.cell.2011.11.033

How does RNA polymerase recognize a promoter in duplex DNA? How are the DNA strands pried apart to enable RNA synthesis? A crystal structure by Feklistov and Darst unexpectedly reveals that these two processes are interconnected.

In a landmark paper on the initiation of transcription, Feklistov and Darst (2011) present in this issue the near atomic resolution structure of a bacterial RNA polymerase fragment bound to a bacterial promoter DNA sequence (Feklistov and Darst, 2011). In so doing, the authors address the central mysteries of transcription initiation: the basis for sequence-specific promoter recognition and the mechanism of DNA melting to expose a single strand to direct RNA synthesis. The remarkable conclusion is that promoter recognition and DNA melting are one and the same: they are coupled; they represent two facets of the same process.

The sequence element studied by Feklistov and Darst occurs 10 bp upstream from the transcription start sites of most bacterial promoters and is bound by region 2 of σ, the subunit of bacterial RNA polymerase responsible for promoter-specific transcription. The "−10" element lies at the edge of an initial melted region (transcription bubble) that extends to the transcription start site at +1, and binding of σ to the nontranscribed strand of the −10 element has been thought to play a role in bubble formation or maintenance (Murakami and Darst, 2003). Indeed, as shown by Feklistov and Darst, σ region 2 binds specifically to a −10 nontranscribed single strand and not to −10 duplex DNA. It is the complex of region 2 with the nontranscribed strand that is crystallized and solved by Feklistov and Darst.

The resulting structure shows a lock-and-key fit of region 2 to the nontranscribed strand (Figure 1). Every nucleotide of the consensus −10 sequence, 5′-$T_{-12}A_{-11}T_{-10}A_{-9}A_{-8}T_{-7}$-3′, interacts with the protein. A_{-11} and T_{-7} are flipped out of the base stack, with their bases buried in pockets on the protein surface. Only an A base can fit in the A_{-11} pocket, and the T base is specifically recognized in the T_{-7} pocket. The intervening bases of $T_{-10}A_{-9}A_{-8}$ remain stacked, directed away from the protein surface; these three nucleotides interact with the protein through their sugar-phosphate backbones. The structure thus accounts for the near absolute conservation of A at position −11 and T at −7 and lesser conservation of the intervening bases.

Recognition of the −10 element has been suggested to begin with the DNA in double-stranded form, but the structure revealed by Feklistov and Darst requires melting to single strands. The authors therefore investigated the binding of region 2 to −10 DNA at 4°C, where melting is strongly disfavored. By inserting base derivatives that disrupt protein single-stranded DNA binding but preserve protein double-stranded DNA interaction, or the reverse, Feklistov and Darst establish that all recognition occurs with the nontranscribed strand in the manner shown by their structure. Specific recognition of the −10 element in duplex form is nearly excluded.

Thus the −10 element must melt for binding, whereas at the same time the element must be bound to preserve melting. Feklistov and Darst suggest that the resolution of this paradox is a sequential process, initiated by flipping of A_{-11} to begin strand separation, followed by unzipping of the rest of the −10 duplex, to arrive at a transcription bubble within which resides the structure reported here. Residues of σ region 2 may serve as "wedges" to facilitate base flipping, as suggested for other protein-DNA interactions. It is important to bear in mind, however, that melting is thermally driven. σ can only supply binding energy to stabilize a state arising through a thermal fluctuation. There is no active (ATP- or other energy-dependent) process.

Recognition and melting are therefore closely coupled. They occur in lockstep fashion. Neither can develop or be sustained without the other. When complete, the bubble with tightly bound σ is exceedingly stable, persisting through multiple rounds of abortive initiation before productive transcription ensues. A parallel may be drawn with eukaryotic transcription, in which stability conferred by interactions with polymerase-associated proteins, the so-called general transcription factors, sustains a transcription bubble during abortive initiation. The outcome is the same: genuine promoter complexes, stabilized by protein-DNA interactions, survive, whereas nonspecific complexes, lacking stabilization, are short lived and dissociate. Promoter specificity is achieved through kinetic proofreading (Liu et al., 2011).

How far does the parallel with eukaryotic transcription extend? Crystallography has revealed remarkable structural homology between one of the general transcription factors, TFIIB, and σ protein (Kostrewa et al., 2009; Liu et al., 2010). The homology is most notable between the C-terminal region of TFIIB and σ region 3, and between the so-called B-linker and B-finger and the σ region 3–region 4 linker (Bushnell et al., 2004; Kostrewa et al., 2009; Liu et al., 2010). There is no significant homology to

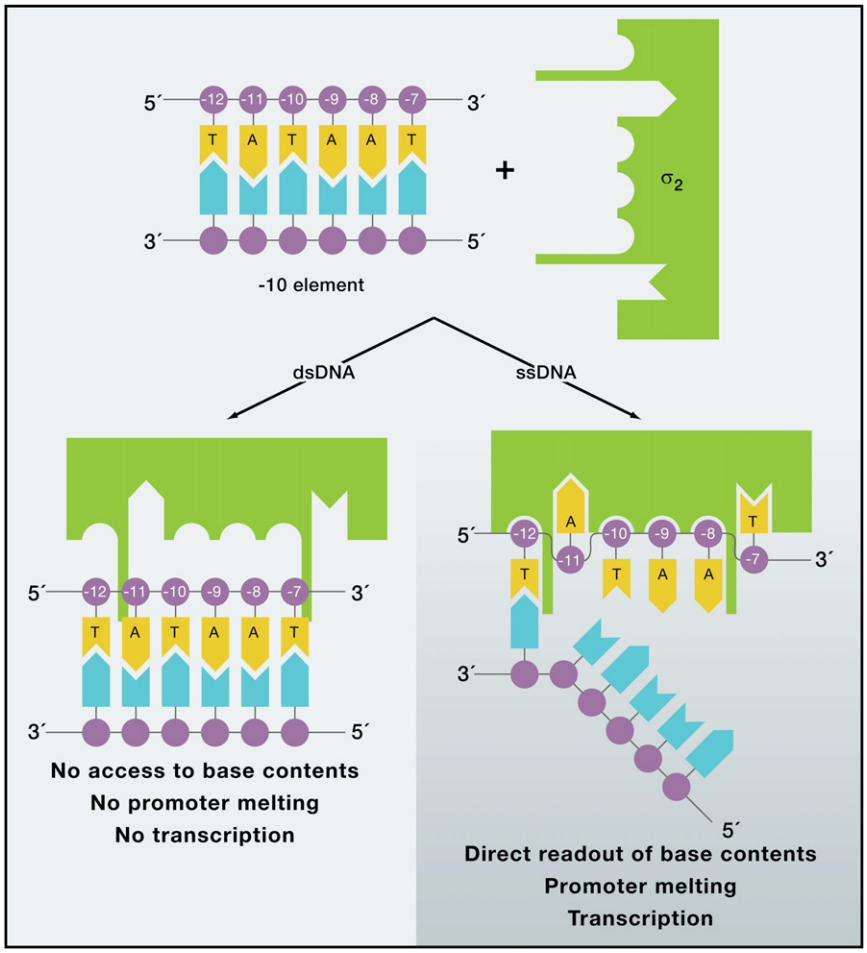

Figure 1. Sequence-Specific Recognition of the −10 Element by Region 2 of σ

The DNA backbone is represented by solid purple circles, bases of the nontemplate strand by yellow polygons, and bases of the template strand by cyan polygons. The sequence of the nontemplate strand corresponds to the consensus of the −10 element, 5′-$T_{-12}A_{-11}T_{-10}A_{-9}A_{-8}T_{-7}$-3′. Region 2 of σ is shown as a green polygon.

σ regions 2 or 4. There is also no eukaryotic promoter sequence as prevalent as the −10 element. Nevertheless, the B-finger projects into the polymerase active center and comes in close proximity to the transcribed DNA strand, with which it may make stabilizing contacts (Bushnell et al., 2004). The possibility has also been raised that a subunit of TFIIF interacts with the nontranscribed strand, performing a similar function to σ region 2.

The formation of a melted promoter through tight specific interaction of the DNA with σ protein begs the question of how the interaction is disrupted to reset the original promoter state. σ-RNA polymerase and σ-promoter interactions are disrupted in stages. The σ region 3–region 4 linker, and region 4 itself, lie in the path of RNA emerging from the polymerase active center and are therefore displaced by RNA synthesis, resulting in the loss of upstream σ-promoter DNA contacts (−35 element; Murakami and Darst, 2003; Nickels et al., 2005). At the same time, synthesis of the initial RNA transcript in the absence of promoter release results in polymerase "scrunching" of the downstream DNA, generating stress that is proposed to facilitate release of σ region 2 contacts with the −10 element (Revyakin et al., 2006; Kapanidis et al., 2006). There is again a parallel with eukaryotic transcription, where the B-finger lies in the path of the RNA and may also be displaced by RNA synthesis. It is still unclear how eukaryotic transcription factor–DNA complexes are disrupted. Mot1 protein may play a role in this process (Auble et al., 1994). Additional such factors or machinery in bacteria remain to be identified.

REFERENCES

Auble, D.T., Hansen, K.E., Mueller, C.G., Lane, W.S., Thorner, J., and Hahn, S. (1994). Genes Dev. *8*, 1920–1934.

Bushnell, D.A., Westover, K.D., Davis, R.E., and Kornberg, R.D. (2004). Science *303*, 983–988.

Feklistov, A., and Darst, S.A. (2011). Cell *147*, this issue, 1257–1269.

Kapanidis, A.N., Margeat, E., Ho, S.O., Kortkhonjia, E., Weiss, S., and Ebright, R.H. (2006). Science *314*, 1144–1147.

Kostrewa, D., Zeller, M.E., Armache, K.J., Seizl, M., Leike, K., Thomm, M., and Cramer, P. (2009). Nature *462*, 323–330.

Liu, X., Bushnell, D.A., Wang, D., Calero, G., and Kornberg, R.D. (2010). Science *327*, 206–209.

Liu, X., Bushnell, D.A., Silva, D.A., Huang, X., and Kornberg, R.D. (2011). Science *333*, 633–637.

Murakami, K.S., and Darst, S.A. (2003). Curr. Opin. Struct. Biol. *13*, 31–39.

Nickels, B.E., Garrity, S.J., Mekler, V., Minakhin, L., Severinov, K., Ebright, R.H., and Hochschild, A. (2005). Proc. Natl. Acad. Sci. USA *102*, 4488–4493.

Revyakin, A., Liu, C., Ebright, R.H., and Strick, T.R. (2006). Science *314*, 1139–1143.